Inertia and Eigenvalue Relations between Symmetrized and Symmetrizing Matrices for the Real and the General Field Case*

Frank Uhlig

Institut für Geometrie und Praktische Mathematik
RWTH Aachen
Templergraben 55
51 Aachen, West Germany
and
Department of Mathematics
Oregon State University
Corvallis, Oregon 97331

Submitted by Hans Schneider

ABSTRACT

Starting from a theorem of Frobenius that every $n \times n$ matrix is the product of two symmetric ones, we study relations between the similarity invariants of a square matrix and the congruence invariants of its symmetric factors. Section 1 treats the real case, Sec. 2 the arbitrary field case, and Sec. 3 the indefinite inner product case for Krein spaces. The proofs are obtained from the real canonical pair form in Secs. 1 and 3 and from the recently found rational canonical pair form in Sec. 2, each time via combinatorial type arguments on weighted partitions of n. The resulting theorems typically give bounds for the elementary divisor structure of A in terms of the index or signature of one or both of its symmetric factors (or vice versa). Our results greatly extend and generalize the classic results of Klein, Loewy, Taussky, et al., and in Sec. 2 put new light on Waterhouse's recent characterization of hereditarily euclidean fields. A short survey on the history of the subject from the early 1800s on completes the paper.

INTRODUCTION

By a theorem of Frobenius [18] every finite dimensional square matrix over an arbitrary field can be expressed as the product of two symmetric

*This paper extends Sec. 7 of my Habilitationsschrift at the University of Würzburg, W. Germany, 1976.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/81/010203+24$02.50

matrices, one of which can be chosen nonsingular. Whence the question arises: how are the similarity invariants of an arbitrary square matrix A (i.e. its eigenvalues, elementary divisors, etc.) related to the congruence invariants of its symmetric factors (i.e. to their inertia, signature, index etc.)?

We will give a description of these relations in Part A, first in Sec. 1 for real matrices, then in Sec. 2 for the arbitrary field case, and finally in Sec. 3 for the infinite dimensional case. Our basic tool in finite dimensions will be the canonical pair form theorems. This approach to the problem can be traced back to various authors in the second half of the last century, when a real canonical pair form was attempted by several authors, but before such a pair form was actually completely understood. Hence the results on the relations among invariants of symmetrized and symmetrizing matrices from the last century ought to be checked. It will turn out that, in fact, they can be significantly extended for both the real and the abstract field case.

When the problems associated with the real canonical pair form were finally fully and correctly understood by Muth [39] in 1905, and the proof thereafter simplified by Trott [51] in 1934 and by Ingraham and Wegner [24] in 1935, there was apparently no longer any interest in our subject. Around 1960, when Lyapunov's inertia theorem had been generalized, largely by Taussky [48, 49] and by Ostrowski and Schneider [42], parts of the old results together with new ones were obtained by Carlson [7, 8] using Lyapunov type methods. Most recently, since the rational canonical pair form theorem became available due to Uhlig [56, 58] and Waterhouse [60], new and amazing results have emerged for the arbitrary field case as well.

Questions like ours have also come up for infinite dimensional spaces from quantum field theory and related mathematical research on indefinite inner product spaces. In Sec. 3 we will show that the results relating to our subject for infinite dimensional spaces are in fact consistent with and easily obtainable from the results for the finite dimensional real case in Sec. 1, provided one interprets them for finite dimensions only. Amazingly enough, an infinite dimensional analogue to Frobenius's theorem seems to be neither known nor disproved, so that the proofs for the infinite dimensional case (as e.g. in Bognar [2]) use different techniques.

We conclude this paper with a historical Part B on the achievements of authors before 1900 and from around 1960 on; there were no papers published in this area in between.

A. MATHEMATICS

0. *Preliminaries*

When studying commuting matrices, Frobenius [18] in 1910 came upon this theorem, which he called "remarkable" (his introduction to it on p. 42

reads: "So ergibt sich der merkwürdige Satz:"):

THEOREM 0. *For every* $A \in F_{nn}$ *there exist* $S = S', T = T' \in F_{nn}$ *such that* $A = ST$ *and* S *or* T *can always be chosen nonsingular.*

Here and throughout this paper T' means the *transpose* of T. This deep algebraic result has influenced the development of matrix theory and algebra less than might be expected. Only 49 years later was it extended significantly by Taussky and Zassenhaus [46, Theorem 2]. In the literature, Theorem 0 is sometimes attributed to Voss [59], but Voss, studying "orthogonal substitutions" alone, proved only a special case of Frobenius's result, namely, that an orthogonal matrix can be expressed as the product of two symmetric ones (see [59, p. 343]). (The statement regarding Voss in MacDuffee [38, p. 80, 4th paragraph] unfortunately is incorrect.) Further references to the factorization theorem can be found in Taussky [50].

We now define some of the concepts that we will use later on.

DEFINITION.

(a) For $A \in F_{nn}$ we set $S_{sn}(A) := \{S \in F_{nn} \mid S = S' \text{ nonsingular with } SA = A'S\}$, the *symmetric nonsingular symmetrizer* of A.

(b) For $S = S' \in F_{nn}$ nonsingular, let $J(S) := \{A \in F_{nn} \mid SA = A'S\}$ be the set of matrices symmetrized by S.

(c) For $T = T' \in R_{nn}$ let $\operatorname{in} T := (i, j, l)$, with $i + j + l = n$, denote the *inertia triple* of T if T has i positive, j negative, and l zero eigenvalues.

(d) For $T = T' \in R_{nn}$ nonsingular, let $\operatorname{sig} T := |i - j|$ be the *signature* of T ($l = 0$ in this case).

(e) For $A \in R_{nn}$ let $u_j(A) := \#(\text{odd dimensional Jordan blocks in the real Jordan normal form of } A)$.

(f) For $T = T' \in F_{nn}$ let $\operatorname{index} T := \max\{\dim V \mid V \subseteq F^n \text{ subspace, } v'Tv = 0 \text{ for all } v \in V\}$.

With this notation we cite a result due to A. Loewy [37, p. 67]:

LEMMA. *If* $S = S' \in R_{nn}$ *is nonsingular, then* $\operatorname{sig} S + 2 \operatorname{index} S = n$.

1. *The Real Case*

The investigations here are based on the real canonical pair form theorem in a simplified version of Uhlig [52, 55]:

REAL CANONICAL PAIR FORM. *For* $S = S', T = T' \in R_{nn}$, S *nonsingular, let* $J = \operatorname{diag}(J_i)$ *be the real Jordan normal form of* $S^{-T}T$. *Then* S *and* T *are* *simultaneously R-congruent to*

$$\operatorname{diag}(\varepsilon_i E_i), \quad \operatorname{diag}(\varepsilon_i E_i J_i), \tag{*}$$

where $\varepsilon_i = \pm 1$, *and* E_i *has the form*

$$\begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix}$$

with $\dim E_i = \dim J_i$.

The simultaneous reduction of real symmetric matrices was originated in the late 1860s by both Kronecker [32] and Weierstrass [63]. For a short history of the real pair form see the introduction in Uhlig [55]. There, more results are proved regarding the ε_i and their uniqueness. But the full statement of the real canonical pair form theorem will not be needed in the following, nor was it necessary for the results obtained before 1900, to be summarized in Part B.

Note that if S and T are R-congruent to (*), then $S^{-1}T$ has real Jordan normal form $J = \operatorname{diag}(J_i)$.

Our first result was obtained in Uhlig [54, Theorem 2]:

THEOREM 1. *If* $A \in R_{nn}$, *then for all* $S \in S_{sn}(A)$,

$$\left.\begin{matrix} 0 & n \text{ even} \\ 1 & n \text{ odd} \end{matrix}\right\} \leqslant \operatorname{sig} S \leqslant u_i(A).$$

Moreover there exist $S \in S_{sn}(A)$ *satisfying*

$$\operatorname{sig} S = u_i(A) \text{ and } \operatorname{sig} S = \left\{ \begin{matrix} 0 \\ 1 \end{matrix} \right.,$$

respectively. All other signatures in between occur in steps of two for $S \in S_{sn}(A)$.

As a corollary one has this variant of a theorem of Drazin and Haynsworth [17]) for nonsingular S (see Uhlig [54, Corollary 2]):

COROLLARY 1.

(a) *If* $\operatorname{sig} S = k$ *for an* $S \in S_{sn}(A)$, *then* A *has at least* k *linearly independent eigenvectors corresponding to real eigenvalues of* A.

(b) *If A has exactly k linearly independent eigenvectors for real eigenvalues, then* $\operatorname{sig} S \leqslant k$ *for each* $S \in S_{sn}(A)$.

Proof. Every odd dimensional Jordan block in the real Jordan normal form of A must correspond to a real eigenvalue of A. Hence $u_j(A) \leqslant k$, and by Theorem 1, $\operatorname{sig} S \leqslant u_j(A) \leqslant k$. ∎

Having thus determined the bounds that a fixed matrix $A \in R_{nn}$ imposes on the inertia of its nonsingular symmetrizers, we will reverse the viewpoint now: how does a given nonsingular $S = S' \in R_{nn}$ determine the structure of matrices $A \in J(S)$ that are symmetrized by it? The answer is the following:

THEOREM 2. *If $S = S' \in R_{nn}$ is nonsingular, then all $A \in J(S)$ satisfy* $\operatorname{sig} S \leqslant u_j(A) \leqslant n$. *All values occur in steps of two between* $\operatorname{sig} S$ *and n for $u_j(A)$ and some $A \in J(S)$.*

Proof. By definition, if $A \in J(S)$, then $SA = A'S$ is symmetric. We are thus led to consider the real canonical pair form for the pair S and SA: its block sizes are determined by the block sizes $\dim J_i$ of the real Jordan normal form $J = \operatorname{diag}(J_i)$ of $S^{-1}SA = A$. And S and SA are simultaneously R-congruent to $\operatorname{diag}(\varepsilon_i E_i)$ and $\operatorname{diag}(\varepsilon_i E_i J_i)$ with $\varepsilon_i = \pm 1$. For a given nonsingular $S = S' \in R_{nn}$, only certain Jordan structures are allowed for an $A \in J(S)$, since by Sylvester's law,

$$\operatorname{sig} S = \operatorname{sig}(\operatorname{diag}(\varepsilon_i E_i)) = \sum_{\substack{\dim E_i \text{ odd} \\ \varepsilon_i = 1}} 1 - \sum_{\substack{\dim E_i \text{ odd} \\ \varepsilon_i = -1}} 1.$$

Hence $\operatorname{sig} S \leqslant \#(\text{odd dimensional blocks } E_i) = u_j(A)$. By Loewy's lemma $\operatorname{sig} S = n - 2 \operatorname{index} S$, so that $\operatorname{sig} S$ always differs from n by an even number. The same clearly holds for $u_j(A)$.

Let $\operatorname{sig} S = m < n$. In order to complete the proof we have to exhibit matrices $A \in J(S)$ with $u_j(A) = m, m+2, \ldots, n-2, n$. To obtain A with $u_j(A) = m$, we set $(1) = \varepsilon_i E_i = J_i$ for $i = 1, \ldots, m$ and

$$\varepsilon_n E_n = \begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix}, \qquad J_n = \begin{bmatrix} 1 & 1 & & & 0 \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ 0 & & & & 1 \end{bmatrix}$$

with $\dim E_n = \dim J_n = n - m$ even. Clearly $\operatorname{sig} S = m = \operatorname{sig}(\operatorname{diag}(\varepsilon_i E_i))$, and by

Sylvester's law there is a nonsingular $X \in R_{nn}$ such that $S = \pm X' \operatorname{diag}(\varepsilon_i E_i) X$. For $A := \pm X^{-1} \operatorname{diag}(J_i) X$ we have $SA = X' \operatorname{diag}(\varepsilon_i E_i) X X^{-1} \operatorname{diag}(J_i) X = X' \operatorname{diag}(\varepsilon_i E_i J_i) X = A'S$, since $E_i J_i = J'_i E_i$. Thus $A \in J(S)$ and clearly $u_j(A) = m = \operatorname{sig} S$. To obtain $A \in J(S)$ with $u_j(A) = m+2$, we split off one block $\varepsilon_{m+1} E_{m+1} = (1)$ from $\varepsilon_n E_n$, for which we use the old name $\varepsilon_n E_n$ again. Note that now $\dim E_n = n-m-1$ is odd. Next set $J_{m+1} = (1)$, J_n as before but of size $n-m-1$, and $\varepsilon_n = -1$. Then $\operatorname{sig}(\operatorname{diag}(\varepsilon_1 E_1, \ldots, \varepsilon_{m+1} E_{m+1}, \varepsilon_n E_n) = m+1 -1 = m = \operatorname{sig} S$, but there are $m+2$ odd dimensional Jordan blocks in $J = \operatorname{diag}(J_i)$ now. And just as before, J is similar to a matrix in $J(S)$. To obtain $A \in J(S)$ with $u_j(A) = m+4 \leqslant n$ one proceeds analogously. Split off two one dimensional blocks from $\varepsilon_n E_n$ (which then has size $n-m-3$, odd), and give them opposite signs: $\varepsilon_{m+2} E_{m+2} = (1) = -\varepsilon_{m+3} E_{m+3}$. And $\pm \operatorname{diag}(\varepsilon_i E_i)$ will still be congruent to S. Next define $J_{m+2} = J_{m+3} = (1)$, and J_n as above with $\dim J_n = \dim E_n = n-m-3$, so that the resulting Jordan form $J = \operatorname{diag}(J_i)$ has $m+4$ odd dimensional blocks and is similar to an $A \in J(S)$. Matrices $A \in J(S)$ with $u_j(A) = m+6, m+8, \ldots, n-2, n$ are constructed similarly. ∎

THEOREM 3. *If $S = S' \in R_{nn}$ is nonsingular, then*

(a) sig $S = n - 2\max_{A \in J(S)}\{\sum_i [n_i/2] \mid n_i = \dim J_i$ *for* $J = \operatorname{diag}(J_i)$, *the real Jordan normal form of* $A\}$, *and*

(b) *for all partitions* $\{n_i\}$ *of* n (*i.e.* $n_i > 0$, $\sum_i n_i = n$) *with* $\sum_i [n_i/2] \leqslant$ index S, *there is* $A \in J(S)$ *with real Jordan normal form* $J = \operatorname{diag}(J_i)$ *and* $n_i = \dim J_i$ *for all* i.

Proof. (a): From A. Loewy's lemma it suffices to show that index $S = \max_{A \in J(S)}\{\sum_i [n_i/2] \mid n_i = \dim J_i$ for $J = \operatorname{diag}(J_i)$, the real Jordan normal form of $A\}$. If $A \in J(S)$ has real Jordan normal form $J = \operatorname{diag}(J_i)$, then S is congruent to $\operatorname{diag}(\varepsilon_i E_i)$ with $\varepsilon_i = \pm 1$, $\dim E_i = \dim J_i$ for all i, and thus index $S = \operatorname{index}(\operatorname{diag}(\varepsilon_i E_i))$. Clearly $\operatorname{index}(\operatorname{diag}(\varepsilon_i E_i)) \geqslant \sum_i \operatorname{index} E_i$. We will show that "=" holds here iff the signs ε_i of all odd dimensional E_i are the same. Note that index $E_i = [(\dim E_i)/2]$. If $\dim E_1$ and $\dim E_2$ are odd while $\varepsilon_1 = -\varepsilon_2$, then by Sylvester's law $\operatorname{diag}(E_0, \varepsilon_3 E_3, \ldots, \varepsilon_m E_m)$ is congruent to diag $(\varepsilon_i E_i)$, where E_0 is chosen so that $\dim E_0 = \dim E_1 + \dim E_2$ is even. And

$$\sum_{i=1}^{n} \operatorname{index} E_i = \sum_{i=1}^{n} \left[\frac{\dim E_i}{2} \right]$$

$$< \frac{\dim E_0}{2} + \sum_{i=3}^{n} \left[\frac{\dim E_i}{2} \right],$$

since

$$\left[\frac{\dim E_1}{2}\right]+\left[\frac{\dim E_2}{2}\right]=\frac{\dim E_1+\dim E_2}{2}-1.$$

Thus index $S \geqslant \Sigma_i [n_i/2]$ if $n_i = \dim J_i$ in the real Jordan normal form of $A \in J(S)$, and index $S = \max_{A \in J(S)} \Sigma_i [n_i/2]$.

(b): The value of $\Sigma_i [n_i/2]$ decreases from its maximum if in the partitioning $\{n_i\}$ of n the number of odd n_i is increased. To increase the number of odd n_i in the partitioning $\{n_i\}$, one even n_i has to be split into two odd numbers or two odd numbers have to be split off from an odd $n_i \geqslant 3$. If one gives the associated new odd dimensional blocks E_i opposite signs ε_i, one does not change the inertia of $\operatorname{diag}(\varepsilon_i E_i)$, which thus remains congruent to S. As has been shown in the proof of Theorem 2, one can then construct an $A \in J(S)$ with Jordan block structure conformal with $\{n_i\}$. ∎

Remark. If in $S = (i, j, 0)$, then clearly index $S = \min(i, j)$ and sig $S = |i - j|$. But for any two nonnegative integers $|i-j| = i + j - 2\min(i, j)$; hence Loewy's Lemma follows.

Theorems 2 and 3 lead to a generalization of an old result for definite matrices from before 1850 and of its extension by Taussky [47]:

Corollary 2. *Let* $S = S' \in R_{nn}$.

(a) *S is definite iff every $A \in J(S)$ is similar to a real diagonal matrix.*

(b) *If S is nonsingular, then* $\operatorname{sig} S \leqslant 1$ *iff for every partitioning $\{n_i\}$ of n there is $A \in J(S)$ with real Jordan normal form* $\operatorname{diag}(J_i)$ *and* $\dim J_i = n_i$ *for all i.*

Proof. (a): Assume $n = \operatorname{sig} S$. By Theorem 2 all $A \in J(S)$ must satisfy $n = \operatorname{sig} S \leqslant u_j(A) \leqslant n$. Thus all $A \in J(S)$ must be real diagonable. Conversely if all $A \in J(S)$ are real diagonable, then $\dim J_i = n_i = 1$ for all their real Jordan blocks J_i. And from Theorem 3(a), $\operatorname{sig} S = n - 2 \times 0 = n$, i.e., S is definite.

(b): From Theorem 3(a), $1 \geqslant \operatorname{sig} S = n - 2\max\{\Sigma_i [n_i/2]\}$; thus $\max\{\Sigma_i [n_i/2]\} \geqslant [n/2]$, so that by Theorem 3(b) all partitions $\{n_i\}$ of n occur for the real Jordan structure constants $n_i = \dim J_i$ for some $S \in J(S)$. If conversely all $\{n_i\}$ occur as Jordan constants for $A \in J(S)$, then setting $n_1 := n$, we have $\operatorname{sig} S \leqslant n - 2[n/2] \leqslant 1$ by Theorem 3(a). ∎

Note that the corollary could have been stated with $[n/2]$ subparts, for if $\operatorname{sig} S = n - 2$, then every $A \in J(S)$ can have Jordan blocks of size 2 at most; if sig $S=n-4$, then Jordan blocks of sizes up to 4 are allowed; and so forth. The converse is also true.

Next we indicate how Theorems 1, 2, 3 and Corollary 2 extend Taussky's theorem from [47], [50]:

THEOREM (T). *For $A \in R_{nn}$ there exists $P \in S_{sn}(A)$ positive definite iff A is similar to a real diagonal matrix.*

The "if" part was discovered by Taussky; the "only if" part follows from a result of Cauchy [12] that a positive definite matrix P and any symmetric matrix (PA in this case) can always be diagonalized simultaneously over the reals. We now give a proof of Theorem (T) based on our previous results:

Proof. Clearly $P \in S_{sn}(A)$ iff $A \in J(P)$. If P is positive definite, then by Corollary 2(a), A must be R-diagonable. Conversely, if $u_j(A)=n$, then by Theorem 1 there is always an $S \in S_{sn}(A)$ with $\operatorname{sig} S=n$, i.e., a definite symmetrizer exists for diagonable A. ∎

Thus far we have dealt only with relations between the invariants of an arbitrary matrix and those of one of its symmetric factors. Due to Frobenius's theorem it appears advisable to study invariance relations between a matrix A and both of its symmetric factors S and $S^{-1}A$. Instead of the signature, the index will play the key role here, since A and thus $S^{-1}A$ may very well be singular. We will study the questions: given A of a certain Jordan structure, what index can $S \in S_{sn}(A)$ and $S^{-1}A$ have, and conversely, given S and T with their respective indices, what real Jordan normal forms can $S^{-1}T$ have?

Initially we need to compute index(diag($\varepsilon_i E_i$)) and index(diag($\varepsilon_i E_i J_i$)) explicitly:

PROPOSITION. *Let S and T have real canonical pair form*

$$\operatorname{diag}(\varepsilon_i E_i), \quad \operatorname{diag}(\varepsilon_i E_i J_i). \tag{*}$$

Assume further that

$J_1,\ldots,J_c$ correspond to complex eigenvalues of $S^{-1}T$, $c \geqslant 0$,

$J_{c+1},\ldots,J_{c+r_e}$ are the even dimensional Jordan blocks for nonzero real eigenvalues of $S^{-1}T$, $r_e \geqslant 0$,

$J_{c+r_e+1},\ldots,J_{c+r_e+r_u}$ are the odd dimensional Jordan blocks for nonzero real eigenvalues, $r_u \geqslant 0$,

$J_{c+r_e+r_u+1},\ldots,J_{c+r_e+r_u+r_{0u}}$ are the odd dimensional Jordan blocks for the eigenvalue zero, $r_{0u} \geqslant 0$, and

$J_{c+r_e+r_u+r_{0u}+1},\ldots,J_{c+r_e+r_u+r_{0u}+r_{0e}}$ are the even dimensional Jordan blocks for the eigenvalue zero, $r_{0e} \geqslant 0$.

Then

$$\operatorname{index} S = \sum_i \left[\frac{\dim J_i}{2} \right] + \min_{c+r_e<i\leqslant c+r_e+r_u+r_{0u}} \{ \#\{\varepsilon_i = 1\}, \#\{\varepsilon_i = -1\} \}$$

and

$$\operatorname{index} T = \sum_i \left[\frac{\dim J_i}{2} \right] + r_{0u} + \min_{\substack{c+r_e<i\leqslant c+r_e+r_u \\ c+r_e+r_u+r_{0u}<j\leqslant c+r_e+r_u+r_{0u}+r_{0e}}}$$

$$\{ \#\{\varepsilon_i\lambda_i>0\} + \#\{\varepsilon_j = 1\}, \#\{\varepsilon_i\lambda_i<0\} + \#\{\varepsilon_j = -1\} \}.$$

Proof. We study $S = \operatorname{diag}(\varepsilon_i E_i)$ first. Note that for a single block E, $\operatorname{index} E = [(\dim E)/2]$. And for each pair of odd dimensional blocks E_i, E_j with opposite signs $\varepsilon_i \varepsilon_j = -1$, there is one additional isotropic line belonging to the maximal isotropic subspace of $\operatorname{diag}(\varepsilon_i E_i, \varepsilon_j E_j)$ besides the $[(\dim E_i)/2] + [(\dim E_j)/2]$ dimensional isotropic subspace just mentioned. Since there are $\min_{\dim J_i \text{ odd}} \{ \#\{\varepsilon_i = 1\}, \#\{\varepsilon_i = -1\} \}$ such pairs of blocks E_i, E_j, the formula for index S follows from the definition of r_u and r_{0u}.

With $T = \operatorname{diag}(\varepsilon_i E_i J_i)$ we have, for one block EJ alone,

$$\operatorname{index} EJ = \operatorname{index} E + \begin{cases} \dim\ker J & \text{if } \dim E \text{ is odd,} \\ 0 & \text{otherwise,} \end{cases}$$

with $\dim\ker J \leqslant 1$. Again one extra isotropic line can be found in the maximal isotropic subspace of T for each pair of blocks $E_i J_i$, $E_j J_j$ of appropriate size with elements of opposite sign on the counter- or cocounterdiagonal. Regarding size, $E_i J_i$, $E_j J_j$ must be either one of the r_u odd dimensional blocks for nonzero eigenvalues of A (with counterdiagonal element $\varepsilon_i\lambda_i$, $\varepsilon_j\lambda_j$ respectively) or one of the r_{0e} even dimensional blocks for the eigenvalue zero (with cocounterdiagonal element ε_i, ε_j respectively), which proves the formula for index T. ■

With this computation we can give a bound for the index difference of the two symmetric factors of a given matrix A:

THEOREM 4. *For $A \in R_{nn}$ every pair $S \in S_{sn}(A)$ and $T = T' \in R_{nn}$ with $S^{-1}T = A$ satisfies*

$$\operatorname{index} T - \operatorname{index} S \leqslant \left[\frac{u_j(A) + \dim\ker A}{2} \right].$$

Proof. Without loss of generality we may assume that S and T are in canonical pair form (*). The ε_i can be chosen arbitrarily here and still $S^{-1}T=\operatorname{diag}(J_i)=A$. To find an upper bound for the index difference we first maximize index T. If λ_+ of the real eigenvalues corresponding to odd dimensional Jordan blocks J_i are positive and λ_- are negative, then $\lambda_+ + a$ values of the r_u numbers $\varepsilon_i\lambda_i$ will be positive, while $\lambda_- - a$ will be negative, where a is an integer which is determined by the actual choice of ε_i for the r_u odd dimensional Jordan blocks for nonzero eigenvalues of A. Let x denote the number of positive ε_i associated with the r_{0e} even dimensional Jordan blocks for the zero eigenvalue. To maximize index T means to maximize $\min\{\lambda_+ + a + x, \lambda_- - a + r_{0e} - x\}$. Plotted as functions of x, the two quantities involved represent an increasing and a decreasing line. Thus the minimum will be maximized where the two lines intersect: $\lambda_+ + a + x = \lambda_- - a + r_{0e} - x$, which yields $x = \lambda_-/2 - \lambda_+/2 - a + r_{0e}/2$. Thus the largest minimum is $\lambda_+ + a + x = [(r_u + r_{0e}/2]$ in the integers. Note here that the maximum is independent of a, i.e. independent of the actual choice of the ε_i. Thus one can choose a so that index S is minimized: As shown in the proof of Theorem 3(a), the index of S is minimal if all ε_i for odd dimensional blocks E_i have the same sign. And then index $S=\Sigma_i[(\dim J_i)/2]$. The result follows finally from the Proposition, namely

$$\operatorname{index} T - \operatorname{index} S \leqslant r_{0u} + \left[\frac{r_u + r_{0e}}{2}\right]$$

$$= \left[\frac{2r_{0u}}{2} + \frac{r_u + r_{0e}}{2}\right] = \left[\frac{u_f(A) + \dim\ker A}{2}\right]. \qquad \blacksquare$$

An easy consequence is

COROLLARY 3. *If $A \in R_{nn}$ is nonsingular, then the indices of a pair of its symmetric factors differ by $[u_f(A)/2]$ at most, while their signatures differ by $2[u_f(A)/2]$ at most.*

Along similar lines one can prove the "converse":

THEOREM 5. *Let $S=S'\in R_{nn}$ be nonsingular and $T=T'\in R_{nn}$. Then $S^{-1}T$ satisfies*

$$\left[\frac{u_f(S^{-1}T) + \dim\ker T}{2}\right] \geqslant \operatorname{index} T - \operatorname{index} S.$$

2. *The Arbitrary Field Case*

We are interested here in invariance relations between symmetrized and symmetrizing matrices over arbitrary fields not of characteristic two. A question that arose naturally out of the results for the real case was raised by Uhlig in [56, Chapter 6, p. 40]: For which fields does Taussky's Theorem (T) hold? Or in other words, classify all fields F for which every $A \in F_{nn}$ that has an anisotropic symmetrizer is F-similar to a diagonal matrix. Here we use the

DEFINITION. $S = S' \in F_{nn}$ is *anisotropic* if $x'Sx = 0$ implies $x = 0$.

This question was partially answered by Waterhouse [61]. Other results for arbitrary fields were obtained simultaneously by Uhlig [56] and Waterhouse [61]. The proofs of these results generally rely on the rational canonical pair form theorem as developed by Uhlig [56, 58] and the slightly different form due to Waterhouse [60].

RATIONAL CANONICAL PAIR FORM. *For $S = S'$, $T = T' \in F_{nn}$, S nonsingular, let* $\operatorname{diag}(Q_i)$ *be the Jacobson normal form of $S^{-1}T$. Then S and T are simultaneously F-congruent to*

$$\operatorname{diag}(S_i), \quad \operatorname{diag}(S_iQ_i). \tag{**}$$

If

$$Q_k = \begin{bmatrix} P_k & & & & 0 \\ N & \cdot & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ 0 & & & N & P_k \end{bmatrix}$$

is the Jacobson matrix for the elementary divisor $p_k^{t_k}$ of $S^{-1}T$, P_k the companion matrix for p_k,

$$N = \begin{pmatrix} & 0 \\ 1 & \end{pmatrix}$$

with $\dim P_k = \dim N = \deg p_k$, $\operatorname{blockdim} Q_k = t_k$, *then S_k is partitioned conformally with Q_k, i.e.,* $\operatorname{blockdim} S_k = t_k$, *and $S_k = V_k f_k(Q_k)$ for some $0 \neq f_k \in$*

$F[x]$,

$$V_k = \begin{bmatrix} 0 & & U_k \\ & \cdot^{\cdot^{\cdot}} & \\ U_k & & 0 \end{bmatrix}$$

with blockdim $V_k = t_k$, *where* $U_k = U_k'$ *is determined by* $p_k \in F[x]$ *such that* $V_k Q_k = Q_k' V_k$.

Williamson [64, 65] attempted to prove a rational pair form theorem in 1935, but his reductions were not without mistakes (see e.g. Uhlig [58, p. 62]). Concurrently with the author, Waterhouse [60] developed another version of a rational pair form that differs from ours by not using Jacobson matrices Q_k but companion matrices for the elementary divisor $p_k^{t_k}$ instead. This slightly coarser pair form, obtained via module theory, will lead to other bounds for the index of a symmetrizer than our rational pair form does.

Here again more is known about the f_k, the converse implication, etc. (see e.g. Uhlig [58, Main Theorem]), but again the full rational pair form theorem does not seem to be needed here.

THEOREM 6. *If* $A \in F_{nn}$ *has elementary divisors* $p_i^{t_i}$ *over* F, *then for every* $S \in S_{sn}(A)$ *we have* index $S \geqslant \Sigma_i \deg p_i [t_i/2]$.

This result clearly corresponds to Theorem 1 for the real case, but it is weaker: If for example $t_i = 1$ for all i, then from Theorem 6 index $S \geqslant 0$, while Theorem 1 together with Loewy's Lemma gives

$$\text{index } S \geqslant \frac{n - u_f(A)}{2} = \frac{n - \#\{\deg p_i \text{ odd}\}}{2}.$$

Proof. We use the rational pair form theorem: each $S \in S_{sn}(A)$ is F-congruent to diag(S_i) where

$$S_i = V_i f_i(Q_i) = \begin{bmatrix} * & & U_i f_i(Q_i) \\ & \cdot^{\cdot^{\cdot}} & \\ U_i f_i(Q_i) & & 0 \end{bmatrix}$$

with blockdim $S_i = t_i$, dim $S_i = t_i \deg p_i$. Thus for each i, index $S_i \geqslant \deg p_i [t_i/2]$, since the "last" $[t_i/2] \deg p_i$ unit vectors span an isotropic space for S_i. Clearly then index $S \geqslant \Sigma_i$ index $S_i \geqslant \Sigma_i \deg p_i [t_i/2]$. ∎

REMARK. M. Kneser has pointed out an alternate proof for Theorem 6 from this well-known result, which in itself is one important step towards the proof of the canonical pair form theorems:

(A). *Let V be a finite dimensional vector space over an arbitrary field F with nondegenerate inner product. If $A: V \to V$ is self-adjoint, then V can be written as the orthogonal direct sum of cyclic $F[A]$ modules V_i where $A_i := A|_{V_i}$ has minimal polynomial $p_i^{t_i}$ with p_i F-irreducible.*

Kneser's proof—for which we are very grateful—proceeds as follows: Set $B|_{V_i} = p_i(A_i)^{t_i - [t_i/2]}$. Then B is self-adjoint and $B^2 = 0$. Thus BV_i is isotropic of the proper dimension.

As in the real case, one can reverse the viewpoint:

THEOREM 7. *Let $S = S' \in F_{nn}$ be nonsingular. Then the elementary divisors $p_i^{t_i}$ of every $A \in J(S)$ satisfy $\Sigma_i \deg p_i [t_i/2] \leqslant \text{index } S$.*

The proof can be omitted since it has become routine by now.

Next we generalize one part of Taussky's Theorem (T) to arbitrary fields, namely, the "only if" part, which was actually known to Cauchy [12] in 1829; see also Uhlig [56, Theorem 6, p. 39].

COROLLARY 4. *If $S = S'$ is anisotropic, then every $A \in J(S)$ is F-similar to* diag(P_i) *where each P_i is a companion matrix for an F-irreducible polynomial $p_i \in F[x]$.*

Proof. Assume index $S = 0$. By Theorem 7, index $S \geqslant \Sigma_i \deg p_i [t_i/2] \geqslant 0$ for all elementary divisors $p_i^{t_i}$ of $A \in J(S)$. Thus $t_i = 1$ for all i, i.e., all Jacobson matrices Q_i for $A \in J(S)$ contain but one block P_i, a companion matrix. ■

This was also proved by Waterhouse [61], who additionally showed via Springer's theorem that each p_i must be separable in this case.

It is interesting to note that for $F = R$ one can recapture the "only if" direction of Cauchy in Theorem (T) from Corollary 4, although Theorems 6 and 7 are certainly weaker than the corresponding results for R in Sec. 1.

COROLLARY 5. *If $S = S' \in R_{nn}$ is definite, then every $A \in J(S)$ is R-diagonable.*

Proof. From Corollary 4, each $A \in J(S)$ must be R-similar to diag(P_i), where P_i is a real companion matrix for an R-irreducible polynomial p_i. In R, these polynomials must have $\deg p_i \leqslant 2$. The proof will be complete if we can show that for

$$P=\begin{pmatrix} 0 & 1 \\ -a^2-b^2 & 2a \end{pmatrix},$$

the companion matrix of $(x-\lambda)(x-\bar{\lambda})$ where $\lambda=a+bi\notin R$, every $S\in S_{sn}(P)$ is isotropic. If for

$$S=\begin{pmatrix} x & y \\ y & z \end{pmatrix}$$

the matrix

$$SP=\begin{pmatrix} * & x+2ay \\ -(a^2+b^2)z & * \end{pmatrix}$$

is symmetric, then $x+2ay=-(a^2+b^2)z$, or $x=-2ay-(a^2+b^2)z$. And $\det S=xz-y^2=-2ayz-(a^2+b^2)z^2-y^2=-((y+az)^2+b^2z^2)<0$, since S is nonsingular. Thus S has one positive and one negative eigenvalue and must be indefinite. ■

Note that for arbitrary fields F there is no stronger version than Corollary 4 regarding F-diagonability of matrices with anisotropic symmetrizers:

Remark. In general, S anisotropic does not imply that every $A\in J(S)$ is F-diagonable.

Example. Let

$$F=Q \quad \text{and} \quad S=\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$A=\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\in J(S),$$

since

$$SA=\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

But the characteristic polynomial of A, $f_A(x)=x^2-2$, is irreducible over Q, so that A cannot be Q-diagonalized.

Waterhouse [61, Proposition 6] gave a necessary condition for fields over which Taussky's Theorem (T) holds via his version of a rational pair form in [60] as follows (a sufficient condition does not seem to be known, though):

THEOREM (W1). *If the condition*

(*T* anic.) *For $A \in F_{nn}$ there exists an anisotropic $S \in S_{sn}(A)$ iff A is F-diagonable.*

holds over F, then F is euclidean or $F = F^2$.

Moreover Waterhouse [61, Theorem 8, Proposition 9] obtained inequalities analogous to those of Theorem 6 and 7 from his rational pair form in [60] that hold for exactly two types of fields:

THEOREM (W2).

(a) *If F is hereditarily euclidean, $S = S' \in F_{nn}$ nondegenerate, then every $A \in J(S)$ with elementary divisors $p_i^{t_i}$ over F satisfies* $\Sigma_i [\deg p_i t_i/2] \leqslant \operatorname{index} S$, *and*

(b) *If $S = S' \in F_{nn}$ is nondegenerate and if the elementary divisors $p_i^{t_i}$ of each $A \in J(S)$ satisfy* $\Sigma_i [\deg p_i t_i/2] \leqslant \operatorname{index} S$, *then F is hereditarily euclidean or $F = F^2$.*

Fields F are called *hereditarily euclidean* if they themselves and each formally real finite algebraic extension are *euclidean*, which in turn is defined as formally real with either $\sqrt{a}$ or $\sqrt{-a}$ in F for every $a \in F$. For various equivalent definitions of hereditarily euclidean fields see e.g. Prestel and Ziegler [43]. Waterhouse's proof of Theorem (W2) proceeds via his pair form for F real closed, since by Springer's theorem (see Lam [34, p. 198]), $\operatorname{index} S|_F = \operatorname{index} S|_{\text{real closure of } F}$. It is certainly amazing that by placing the "greatest integer" brackets slightly differently than we have done in Theorem 6 and 7, Waterhouse can classify very specific types of fields, while positioning the square brackets as we did gives results valid for all fields not of characteristic two. An explanation of this phenomenon, unfortunately, does not seem to be known.

Another way of extending Taussky's Theorem (T) would be to work over formally real fields and study fields where matrices with a positive definite symmetrizer are always diagonable:

THEOREM (W3). *If F is formally real, then*

(T pos. def.) *For $A \in F_{nn}$ there exists a positive definite $S \in S_{sn}(A)$ iff A is F-diagonable.*

holds for F iff F is the intersection of its real closures.

This was obtained by Waterhouse [61, Proposition 7].

3. The Infinite Dimensional Case

In this section we will mention several results for indefinite inner product spaces that became available through the book *Indefinite Inner Product Spaces* by Bognar [2] in 1974. The results that we will mention were originated by Iovidov and Krein [25, 26], Glazman and Ljubic [19], and Bognar [1]. For the roles of their individual contributions see the "Notes to Chapter IX" in Bognar [2, pp. 207–209].

Consider an indefinite inner product $[x, y]$ on an infinite dimensional vector space X. An operator A with domain $D_A = X$ is called symmetric with respect to $[\ldots,\ldots]$ if $[x, Ay] = [Ax, y]$ for all $x, y \in X$. In stating the results from Bognar [2], we will assume that a Hilbert space with its definite inner product $(\ldots,\ldots)$ is underlying the indefinite inner product space X with $[\ldots,\ldots]$. This assumption is certainly not necessary for the theory and it is not made in Bognar's book. But when trying to compare the results in [2] with ours for finite dimensions, it is quite natural to do so in the following sense: For $X = R^n$ with the standard positive definite inner product $(x, y) = x'Iy$, we set $[x, y] := x'Sy$ for some nonsingular indefinite $S = S' \in R_{nn}$. If $[x, Ay] = [Ax, y]$, then $x'SAy = (Ax)'Sy = x'A'Sy$ for all $x, y \in R^n$ and thus $SA = A'S$, or $A \in J(S)$ in our notation. The results in [2] that most resemble ours hold for Krein and Pontryagin spaces.

DEFINITION.

(a) $X, [\ldots,\ldots]$ is a *Krein space* if

(1) $[\ldots,\ldots]$ is an indefinite inner product on $X \times X$, i.e., $[\ldots,\ldots]$ is bilinear and $[x, y] = \overline{[y, x]}$,
(2) $X^{\pm} = \left\{ x \in X \mid [x, x] \begin{matrix} \geqslant & 0 \\ \leqslant & 0 \end{matrix} \right\}$ are each complete relative to the "norm" $|x| := |[x, x]|$, and
(3) $X = X^+ \dotplus X^-$.

(b) A Krein space for which $\dim X^+ = k < \infty$ is called a *Pontryagin space* Π_k with rank of positivity k.

(c) For a linear operator T: $X \to X$ for which $T - \lambda I$ is not invertible, we call $S_\lambda(T) = \bigcup_{j=1}^{\infty} \ker((T - \lambda I)^j)$ the *principal subspace* for λ.

(d) An eigenvalue λ of T is *semisimple* if $S_\lambda(T) = \ker(T - \lambda I)$.

These definitions stem from Bognar [2. pp. 29, 100, 184]. The results due to Iovidov and Krein [25, 26], Glazman and Ljubic [19], and Bognar [1] are thus (see Bognar [2, Theorem 4.6, Corollary 4.7, Theorems 4.8, and 4.9, p. 191]):

THEOREM (IKGLB). *Let A be a symmetric operator with respect to the indefinite inner product $[\ldots,\ldots]$ in a Pontryagin space Π_k.*

(a) *If* $\lambda_1,\dots,\lambda_m \notin R$ *are distinct and not complex conjugate eigenvalues of* A, *then* $\sum_{j=1}^{m}\dim_C S_{\lambda_j}(A) \leqslant k$.

(b) *A can have at most k eigenvalues in the half plane* $\operatorname{Im} z > 0$ *and at most k eigenvalues in the half plane* $\operatorname{Im} z < 0$.

(c) *A can have at most k nonsemisimple eigenvalues.*

(d) *The length p of a chain of principal vectors* x_j *with* $Ax_j = \lambda x_j + x_{j-1}$ *for* $j = 0, 1, \dots, p-1$ *where* $x_{-1} := 0$ *and* λ *is a real eigenvalue of* A, *cannot exceed* $2k+1$.

Note that these results were proved without the use of a canonical pair form. Our aim now is to derive all of these results for the finite dimensional real case by use of Sec. 1. Π_k is then to be interpreted as R^n with an indefinite inner product $x'Sx$, where in $S = (k, n-k, 0)$ and $k = \min(k, n-k) = \operatorname{index} S \leqslant [n/2]$. Without loss of generality we will assume S and SA to be in real canonical pair form (*).

Proof. (d): If one $A \varepsilon J(S)$ had real Jordan normal form $\operatorname{diag}(J_i)$, where for one index j, $\dim J_j = m > 2k+1$, then $\operatorname{index}(\varepsilon_j E_j) = [m/2] \geqslant k+1$ and thus $\operatorname{index} S \geqslant k+1$, a contradiction.

(c): If A has $m > k$ nonsemisimple eigenvalues, then A has real Jordan normal form $J = \operatorname{diag}(J_1, \dots, J_m, J_{m+1}, \dots, J_l)$ with $\dim J_i \geqslant 2$ for $i \leqslant m$, while $\dim J_i \geqslant 4$ for complex eigenvalues as long as $i \leqslant m$. For the associated matrix $S = \operatorname{diag}(\varepsilon_i E_i)$ we thus have $\operatorname{index} S = \operatorname{index}(\operatorname{diag} \varepsilon_i E_i) \geqslant m > k = \operatorname{index} S$, a contradiction.

(a): Let A have m pairs of complex conjugate eigenvalues $\lambda_1, \bar\lambda_1, \lambda_2, \bar\lambda_2, \dots, \lambda_m, \bar\lambda_m$ for $\lambda_i \notin R$. Then the real Jordan normal form J of A contains m blocks $J_1, \dots, J_m$ of the form

$$J_j = \begin{pmatrix} a_j & b_j & 1 & 0 & & & & \\ b_j & a_j & 0 & 1 & & & 0 & \\ & & \cdot & & \cdot & & & \\ & & & \cdot & & \cdot & & \\ & & & & \cdot & & \cdot & \\ & & & & & \cdot & 1 & 0 \\ & & & & & \cdot & 0 & 1 \\ & 0 & & & & & a_j & b_j \\ & & & & & & b_j & a_j \end{pmatrix}$$

with $\dim J_j = 2n_j$, $\lambda_j = a_j + ib_j \notin R$ (see e.g. Uhlig [55, Definition 3]). If the pairs $\lambda_j, \bar\lambda_j$ are distinct, we have $\dim_C S_{\lambda_j}(A) = n_j$, while $\operatorname{index}(\varepsilon_j J_j) = n_j$. Thus $\sum_{j=1}^{m}\dim_C S_{\lambda_j}(A) = \sum_{j=1}^{m} n_j = \sum_{j=1}^{m}\operatorname{index}(E_j J_i) \leqslant \operatorname{index} S = k$.

(b) follows from (a), since the complex eigenvalues of a real matrix A occur as pairs of complex conjugate numbers and A can have at most k nonreal, distinct, non complex conjugate eigenvalues. ∎

B. HISTORY

The earliest results relating to this paper go back about 150 years to A. Cauchy [10–13]. Cauchy showed that a real symmetric matrix has only real characteristic roots (see MacDuffee [38], specifically the comments after Corollary 18.31, p. 26), that a symmetric matrix can always be diagonalized by orthogonal congruence [13, Theorem 2], and that a positive definite symmetric matrix can always be diagonalized by congruence simultaneously with any other symmetric matrix [12]. In 1843, Kummer [33] showed that 3×3 symmetric matrices can be orthogonally diagonalized. Cauchy's result (for arbitrary n) was re-proved in yet another way by Borchardt [3], via Vandermonde's determinantal equation, the fact that $\det A^2 = (\det A)^2$ from Jacobi [27, p. 312], and what is now called the Cauchy-Binet formula in Cauchy [10]. In 1853, Sylvester [45, Art. 1] showed that for a definite real symmetric matrix $S=S'$ and $T=T'$, $\det(\lambda S+T)=0$ has only real roots, a result which had already been known to Cauchy, though not in this form. Moreover Sylvester showed in Art. 3 that $\operatorname{sig} S \leqslant \#(\text{real roots of } \det(\lambda S+T))$ and that there exists $T=T'$ such that "$=$" holds. This is the earliest result in the vein of Theorem 1 from Part A. Sylvester's proof is rather sketchy, which prompted A. Ostrowski [40, footnote on p. 3] to describe it as having been done by "*unvollständige Induktion*". In his 1858 paper, Weierstrass [62] credits Cauchy and Jacobi with first studying the roots of $\det(\lambda S+T)=0$ for S, T symmetric. On p. 239 he attributes the following result to Cauchy, Borchardt, and Sylvester: Let $\det S\neq 0$. If $\det(\lambda S+T)=0$ has distinct roots and S or T is definite, then they are all real. On pp. 242, 243 Weierstrass proves this result without assuming that the roots of $\det(\lambda S+T)=0$ are distinct. Hesse [22, 23] and Gundelfinger [20] further elaborated on the results of Kummer, Borchardt, and Weierstrass. Clebsch [15] apparently first studied the eigenvalues of hermitian matrices. On p. 327 he showed that 3×3 hermitian matrices have only real eigenvalues. Later, in [16], he gave a unified proof for arbitrary n that for $A=A'$ or $A=A^*$, $\det(A-\lambda I)=0$ can only hold for real λ. In the same paper it is also shown that a square matrix A with $A=-A'$ or $A=-A^*$ can only have purely imaginary eigenvalues (see p. 235). In 1864 Christoffel [14] generalized the results of Cauchy, Weierstrass, and Clebsch on p. 255: If S and T are hermitian and S is definite, then $\det(\lambda S+T)=0$ has only simple real roots. This result is quoted in MacDuffee [38] as Corollary 36.92 on p. 65 and Theorem 36.8 on p. 64.

Christoffel notes in a footnote on p. 256 that Hermite [21] has already shown that $\det(\lambda S+T)=0$ has only real roots, so that part of Clebsch's results in [15], [16] were known before.

The results mentioned thus far are just slightly connected—one could say embryonically connected—to the main questions of this paper. But by including them here, one obtains a good perspective of the train of thought and the development of matrix and inertia theory in the early 1800s. Many of the results mentioned here were originally not expressed in matrix terminology as we have done, but via determinants, a fact also noted by H. Schneider [44, footnote 2, p. 216; footnotes 13, 14, p. 217].

Felix Klein [31], in his dissertation in 1868, was the first to use canonical pair forms to tackle these problems. On p. 559, Eq. (23), he quotes the pair form for real symmetric matrix pairs of Weierstrass [62], which, unfortunately, is derived via complex congruences. This shortcoming was not noted until 1905, by Muth [39]. But the original "un"-canonical pair form of Weierstrass did not falsify Klein's results. Note that we obtained our results without using the pair form theorems in their most precise form. From Weierstrass's pair form, Klein derives on p. 562 Sylvester's result that $\det(\lambda S+T)=0$ has a least $\operatorname{sig} S$ or $\operatorname{sig} T$ real roots, which is also quoted in MacDuffee [38, Theorem 35.91, p. 65]. On p. 563, Klein obtains the following: if a set of elementary divisors and a symmetric S with $\operatorname{sig} S \leqslant \#$(odd dimensional elementary divisors) are given, then there exists a $T=T'$ such that $S^{-1}T$ has the prescribed elementary divisors—a result which is obvious from the proof of Theorem 2. Thirty years later, in his *Habilitationsschrift*, A. Loewy [35] studied the so-called automorph equation $A^*SA=S$ and related the index of S to the elementary divisor structure of A. Immediately thereafter he noticed the connection to Klein's dissertation. In [36, p. 589] A. Loewy states: For a nonsingular real symmetric matrix pair $(P,Q$, $\operatorname{index} S \geqslant \sum_i[(\dim J_i)/2]$ for all S in the pencil generated by P and Q, where $J=\operatorname{diag}(J_i)$ is the real Jordan normal form of $P^{-1}Q$. Since the inertia of matrices in the pencil generated by P and Q is essentially determined by the Jordan structure of $P^{-1}Q$, A. Loewy's result is contained in Theorem 3(a) and the Proposition. On pp. 590, 591, A. Loewy reestablished Klein's second result in terms of the index. On p. 591 he uses the maximal size m of a Jordan block for matrices in the pencil $aP+bQ$ for variable Q as a means for defining index P, namely $\operatorname{index} P:=[m/2]$, a result reminiscent of Corollary 2(b). On pp. 591, 592 these results are extended to hermitian matrices. None of these results are proved in [36], but rather in [37]. A. Loewy defines $\operatorname{char} T:= \operatorname{index} T-\dim\ker T$ and proves in [37, p. 55] that if S is nonsingular and $S^{-1}T$ has real Jordan normal form $J=\operatorname{diag}(J_i)$, then

$$\operatorname{char} T \geqslant \sum_{\lambda\neq 0}\left[\frac{\dim J_i}{2}\right]+\sum_{\lambda=0}\left[\frac{\dim J_i-1}{2}\right],$$

which is part of our Proposition. An analogous result is proved for nonsingular pencils on pp. 61, 62 via the canonical pair form; see also MacDuffee [38, Theorem 36.9, p. 65]. On pp. 62, 63, A. Loewy re-proves Sylvester's and Klein's result that $\operatorname{sig} S \leqslant \#(\text{real eigenvalues of } S^{-1}T)$. Further on, on p. 67, there are results for hermitian pairs as well. In the same year, T. A. Bromwich [4] re-proved A. Loewy's result that $\operatorname{char} T$ is bounded by the degrees of the elementary divisors of the pencil $aS+bT$ even for singular pencils, by adapting Weierstrass's pair form in [63] to singular pencils. There Bromwich also improves one of Klein's results: If $T=T'$ with $\operatorname{in} T=(k,l,m)$, then $|k-l| \leqslant \#(\text{odd real nonzero elementary divisors of } S^{-1}T)+\#(\text{even dimensional zero elementary divisors of } S^{-1}T)$ for any nonsingular $S=S'$, which follows from our Proposition. These results are re-proved in more detail in Bromwich [5, p. 349].

Summarizing the results obtained in the last century, one notices that they are all contained in and significantly weaker than our results in Sec. A.1.

In 1959, A. Ostrowski [41] studied products of hermitian matrices once again via canonical pair forms. He obtained "Satz 1°" on p. 2: For A, B hermitian with $\operatorname{in} A=(k,l,m)$ and $\operatorname{in} B=(u,v,w)$, $\#(\text{real eigenvalues of } AB) \geqslant \max\{|k-l|+m,\ |u-v|+w\}$. Ostrowski also studied the number of eigenvalues of AB in the right half plane in relation to the inertias of A and B, in Satz 2 and Satz 3. Satz 4 on p. 3 is Sylvester's and Klein's theorem. Further results relating the inertia of AB to the inertias of the hermitian factors A and B were also given by Johnson [66] recently. In 1965, D. Carlson [8] obtained a variety of results by using Lyapunov's equation and stability theory. For example, in Theorem 1, p. 1121, Drazin and Haynsworth's result [17] and Corollary 1 from Sec. 1 are re-proved, as well as Bromwich's extension of the Sylvester-Klein theorem. Corollary 1 on p. 1122 extends this result to hermitian H_1, H_2 thus: $\operatorname{sig} H_1 \leqslant \#(\text{real eigenvalues of } H_1H_2)$, and if H_2 is nonsingular, then $\operatorname{sig} H_2 \leqslant u_j(H_1H_2)$ while $\operatorname{sig} H_1 \leqslant u_j(H_1H_2)+\#(\text{even dimensional Jordan blocks for the zero eigenvalue of } H_1H_2)$. Corollary 3, p. 1124, re-proves Christoffel's result [14] that for definite H_1, H_1H_2 is diagonable with real eigenvalues. Carlson furthermore studies extensions of Frobenius's theorem to products of hermitian matrices. In Theorem 2, p. 1122, he characterizes products of two hermitian matrices as those complex matrices that are similar to a real matrix, while in Theorem 3, p. 1125, hermitian-symmetrizable matrices are studied. See also Carlson [9] and Waterhouse [61, Theorem 2, Proposition 7] for further extensions of Frobenius's theorem in this vein.

It was through my contact with Wallace Givens and Olga Taussky-Todd that I became interested in and familiar with this subject. I am also grateful to have had the advice of D. Carlson, W. D. Geyer, M. Kneser, A. Pfister, and

A. Prestel in the preparation of this paper, an early version of which was read at the 1978 "Quadratische Formen" Conference at Oberwolfach, W. Germany. A large part of the results in Sec. A. 1 and 2 were announced in the survey of definite matrix pencils [57]. Specifically, I wish to thank Linda Hall Library, Kansas City, and its staff for helping me locate several of the pre-1850 articles quoted here. For some of these, only vague references, such as the author and journal volume, were previously available from other 19th century papers.

REFERENCES

1 J. Bognar, On the existence of square roots of an operator which is self-adjoint with respect to an indefinite metric, *Magyar Tud. Akad. Mat. Kutato Int. Közl.* 6: 351–363 (1961) (in Russian).
2 J. Bognar, Indefinite inner product spaces, Ergebnisse der Mathematik Bd. 78, Springer, 1974.
3 C. W. Borchardt, Neue Eigenschaft der Gleichung, mit deren Hülfe man die secularen Störungen der Planeten bestimmt, *J. Reine Angew. Math.* 30: 38–45 (1846).
4 T. J. Bromwich, Note on Weierstrass's reduction of a family of bilinear forms, *Proc. London Math. Soc. Ser. 1* 32, 158–163 (1900).
5 T. J. Bromwich, On a canonical reduction of bilinear forms (part II), with special consideration of congruent reductions, *Proc. London Math. Soc. Ser. 1* 32: 321–352, (1900).
6 D. Carlson and H. Schneider, Inertia theorems for matrices: the semi-definite case, *J. Math. Anal. Appl.* 6, 430–446 (1963).
7 D. Carlson, Rank and inertia bounds for matrices under $R(AH) \geqslant 0$, *J. Math. Anal. Appl.* 10: 100-111 (1965).
8 D. Carlson, On real eigenvalues of complex matrices, *Pacific J. Math.* 15: 1119–1129 (1965).
9 D. Carlson, A note on matrices over extension fields, *Duke Math. J.* 33: 503–506 (1966).
10 A. Cauchy, Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opérées entre les variables qu'elles renferment, *J. École Polytechnique (Paris) X*: 29–112 (1815) (see [3, p. 40]).
11 A. Cauchy, Note sur la polarisation chromatique, *Comptes Rendues* 25: 331–333 (1847) (see[15, p. 319]).
12 A. Cauchy, *Comptes Rendues* 30 (1850) (see [15, p. 320]).
13 A Cauchy, Sur l'équation à l'aide de laquelle on detérmine les inégalités séculaires des mouvements des planètes, *Exercises de Mathématiques, (Paris)* 4: 140–160 (1829); *Oeuvres*, Ser. 2, Vol. 9, pp. 174–195.
14 E. B. Christoffel, Verallgemeinerung einiger Theoreme des Herrn Weierstrass, *J. Reine Angew. Math.* 63: 255–272 (1864).

15 A. Clebsch, Theorie der circularpolarisirenden Medien, *J. Reine Angew. Math.* 57: 319–358 (1860).
16 A. Clebsch, Ueber eine Classe von Gleichungen, welche nur reelle Wurzeln besitzen, *J. Reine Angew. Math.* 62: 232–245 (1863).
17 M. Drazin and E. Haynsworth, Criteria for the reality of matrix eigenvalues, *Math. Z.* 78: 449–452 (1962).
18 F. G. Frobenius, Über die mit einer Matrix vertauschbaren Matrizen (1910), *Ges. Abh.*, Bd. 3, Springer, 1968, pp. 415–427.
19 I. M. Glazman and J. I. Ljubic, *Finite Dimensional Linear Analysis*, Nauka, Moscow, 1969 (in Russian).
20 S. Gundelfinger, Vorlesungen über analytische Geometrie der Ebene, (Dingeldey, Ed.), Leipzig, 1895, pp. 64–74.
21 Hermite, *Comptes Rendus* 41: 181–183 (1855).
22 O. Hesse, Neue Eigenschaften der linearen Substitutionen, welche gegebene homogene Functionen des zweiten Grades in andere transformiren, die nur die Quadrate der Variabeln enthalten, *J. Reine Angew. Math.* 57: 175–182 (1860).
23 O. Hesse, Vorlesungen über analytische Geometrie der Ebene, 3rd ed., Leipzig, p. 515.
24 M. Ingraham and K. Wegner, The equivalence of pairs of hermitian matrices. Trans. Amer. Math. Soc. 38: 145–162 (1935).
25 I. S. Iovidov and M. G. Krein, Spectral theory of operators in spaces with an indefinite metric. I. *Trudy Moskov. Mat. Obšč.* 5: 367–432 (1956); 6: 486 (1957) (in Russian).
26 I. S. Iovidov and M. G. Krein, Spectral theory of operators in spaces with an indefinite metric. II. *Trudy Moskov. Mat. Obšč.* 8: 413–496 (1959); 15: 452–454 (1966) (in Russian).
27 C. Jacobi, De formatione et proprietatibus determinantium, *J. Reine Angew. Math.* 22: 285–318 (1841).
28 C. Jacobi, De functionibus alternantibus earumque divisione per productum e differentiis elementorum conflatum, *J. Reine Angew. Math.* 22: 360–371 (1841).
29 C. Jacobi, Sulla condizione di uguaglianza di due radici dell'equazione cubica, dalla quale dipendono gli assi principali di una superficie del second'ordine, *J. Reine Angew. Math.* 30: 46–50 (1846).
30 C. Jacobi, (same title as [29]), *Giorn. Arcadio* XCIX: 3–11 (1844).
31 F. Klein, Ueber die Transformation der allgemeinen Gleichung des zweiten Grades zwischen Linien-Coordinaten auf eine canonische Form (Dissertation, Bonn, 1868), *Math. Ann.* 23: 539–578 (1884).
32 L. Kronecker, Über Scharen von quadratischen Formen, *Monatsber. Akad, Wiss. Berlin*, 1868, pp. 310–338; *Gesammelte Werke*, Bd. 1, pp. 165–174.
33 E. E. Kummer, Bemerkungen über die cubische Gleichung, durch welche die Haupt-Axen der Flächen zweiten Grades bestimmt werden, *J. Reine Angew. Math.* 26: 268–272 (1843).
34 T. Y. Lam, *The Algebraic Theory of Quadratic Forms*, Benjamin, Reading, Mass., 1973.
35 A. Loewy, Ueber bilineare Formen mit conjugirt imaginären Variablen, *Math. Ann.* 50: 557–576 (1898).

36 A. Loewy, Ueber die Characteristik einer reellen quadratischen Form von nicht verschwindender Determinante, *Math. Ann.* 52: 588–592 (1899).
37 A. Loewy, Ueber Scharen reeller quadratischer und Hermitescher Formen, *J. Reine Angew. Math* 122: 53–72 (1900).
38 C. C. MacDuffee, *The Theory of Matrices*, Ergebnisse der Mathematik, Bd. 2, 2nd ed., Chelsea, New York, 1956.
39 P. Muth, Über reelle Äquivalenz von Scharen reeller quadratischer Formen, *J. Reine Angew. Math.* 128: 302–321 (1905).
40 A. Ostrowski, Über Eigenwerte von Produkten Hermitescher Matrizen, *Hamburger Math. Abh.* 23: 60–68 (1959).
41 A. Ostrowski, Über Produkte Hermitescher Matrizen und Büschel Hermitescher Formen, *Math. Z.* 72: 1–15 (1959).
42 A. Ostrowski and H. Schneider, Some theorems on the inertia of general matrices, *J. Math. Anal. Appl.* 4: 72–84 (1962).
43 A. Prestel and M. Ziegler, Erblich euklidische Körper, *J. Reine Angew. Math.* 274/275: 196–205 (1975).
44 H. Schneider, Olga Taussky-Todd's influence on matrix theory and matrix theorists, a discursive personal tribute, *Linear and Multilinear Algebra* 5: 197–224 (1977).
45 J. J. Sylvester, The algebraical theory of the secular-inequality determinantive equation generalized, *Philos. Mag.* 6: 214–216 (1853).
46 O. Taussky and H. Zassenhaus, On the similarity transformation between a matrix and its transpose, *Pacific J. Math.* 9: 893–896 (1959).
47 O. Taussky, Problem 4846. *Amer. Math. Monthly* 66: 427 (1959); 67: 192, 193 (1960).
48 O. Taussky, A remark on a theorem by Lyapunov, *J. Math. Anal. Appl.* 2: 105–107 (1961).
49 O. Taussky, A generalisation of a theorem by Lyapunov, *J. Soc. Ind. Appl. Math.* 9: 640–643 (1961).
50 O. Taussky, The role of symmetric matrices in the study of general matrices, *Linear Algebra and Appl.* 5: 147–154 (1972).
51 G. Trott, On the canonical form of a nonsingular pencil of hermitean matrices. *Amer. J. Math.* 56: 359–371 (1934).
52 F. Uhlig, A study of the canonical form for a pair of real symmetric matrices and applications to pencils and to pairs of quadratic forms, Thesis, California Institute of Technology, Pasadena, 1971/72.
53 F. Uhlig, Simultaneous block diagonalisation of two real symmetric matrices, *Linear Algebra and Appl.* 7: 281–289 (1973).
54 F. Uhlig, On the nonsingular symmetric factors of a real matrix, *Linear Algebra and Appl.* 8: 351–354 (1974).
55 F. Uhlig, A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil, *Linear Algebra and Appl.* 14: 189–209 (1976).
56 F. Uhlig, A rational pair form for a nonsingular pair of symmetric matrices over an arbitrary field F with char $F \neq 2$ and applications, Habilitationsschrift, Würzburg, 1976.
57 F. Uhlig, A recurring theorem about pairs of quadratic forms and extensions: a

survey, *Linear Algebra and Appl.* 25: 219–237 (1979).
58 F. Uhlig, A rational canonical pair form for a pair of symmetric matrices over an arbitrary field F with char $F \neq 2$ and applications to finest simultaneous block diagonalisations, *Linear and Multilinear Algebra* 8: 41–67 (1979).
59 A. Voss, Zur Theorie der orthogonalen Substitutionen, *Math. Ann.* 13: 320–374 (1878).
60 W. C. Waterhouse, Pairs of quadratic forms, *Invent. Math.* 37: 157–164 (1976).
61 W. C. Waterhouse, Self-adjoint operators and formally real fields, *Duke Math. J.* 43: 237–243 (1976).
62 K. Weierstrass, Über ein die homogenen Functionen zweiten Grades betreffendes Theorem, nebst Anwendung desselben auf die Theorie der kleinen Schwingungen, *Monatsber. Akad. Wiss. Berlin*, 1858, pp. 207–220; *Gesammelte Werke*, pp. 233–246.
63 K. Weierstrass, Zur Theorie der quadratischen und bilinearen Formen, *Monatsber. Akad. Wiss. Berlin*, 1868, pp. 310–338.
64 J. Williamson, The equivalence of a nonsingular pencil of hermitian matrices in an arbitrary field, *Amer. J. Math.* 57: 475–490 (1935).
65 J. Williamson, Note on the equivalence of nonsingular pencils of hermitian matrices, *Bull. Amer. Math. Soc.* 51: 894–897 (1945).
66 C. R. Johnson, The inertia of a product of two hermitian matrices, *J. Math. Anal. Appl.* 57: 85–90 (1977).

Received 3 August 1979; revised 3 March 1980